Combinar les observacions i mesurar la incertesa: Història d'un intent per entendre millor el món*

Marta Sanz i Solé

El naixement de la teoria de la probabilitat s'associa ben freqüentment a l'estudi dels jocs d'atzar. Aquesta activitat lúdica, que segons sembla era àmpliament practicada pels nobles francesos pels volts del segle XVII, és a dir, poc abans de la il·lustració, va plantejar problemes interessants, la majoria de caràcter combinatori, que varen copsar l'atenció de molts científics de l'època. No hi ha dubte que aquest va ser un motor important en el desenvolupament de la probabilitat, però no l'únic.

Al llarg d'aquesta lliçó farem una anàlisi des d'una perspectiva històrica de la gènesi del model lineal de regressió. Com veurem, aquest ha estat un tema que ha plantejat problemes importants que avui formen part de la teoria de la probabilitat i que ha tingut aportacions de matemàtics notables (per exemple, Euler, Legendre, Laplace, Gauss, Cauchy, etc., per citar-ne uns quants).

El model lineal va ser proposat com a solució possible al problema de combinar de manera òptima diferents observacions o amidaments d'un mateix fenomen, realitzats sota condicions idèntiques, per tal de proposar un resultat fiable. Hi ha aquí dos ingredients que cal tenir en compte. D'una banda, l'ingredient estadístic en la pràctica de combinar dades i, d'altra banda, la influència de la física, l'astronomia i la geodèsia en els problemes que van motivar la formulació del model lineal. Ambdós ajuden a posar en evidència dos fets rellevants: la simbiosi estadística-probabilitat, que ha perdurat fins avui, i la intervenció de motivacions procedents d'un cos científic consolidat (en contraposició als jocs d'atzar) en el desenvolupament de la probabilitat.

Resumint, amb el model lineal com a tema principal, intentarem desterrar la idea que la probabilitat és una cosa associada exclusivament als jocs d'atzar, i posarem de manifest les motivacions que varen influir en la seva formulació, els mètodes que es van desenvolupar per solucionar-la i els problemes que es van deixar oberts per a investigacions posteriors.

* Conferència inaugural del curs 1990-1991. 3 d'octubre de 1990. Facultat de Matemàtiques. Universitat de Barcelona.

1. Introducció al model lineal

L'observació de fenòmens físics es fa molt sovint mitjançant aparells de mesura i això comporta necessàriament errors. La qüestió que hem citat abans de com combinar de manera òptima diverses observacions d'un mateix fenomen té com a objectiu que els errors es destrueixin o esdevinguin molt petits.

Al llarg del segle XVIII aquest problema va preocupar obertament científics interessats en qüestions d'astronomia i geodèsia, a causa, ben segur, de la importància de les qüestions que intentaven resoldre. Destaquem-ne les següents:

1. Descriure matemàticament els moviments de la Lluna.
2. Determinar la forma de la Terra.
3. Explicar els moviments de Júpiter i de Saturn.

Les dues primeres tenien clares connotacions comercials i militars, mentre que la tercera estava més en la línia de la pura elucubració científica. Totes tres eren difícils. Com a anècdota, citarem que Newton comentà a Halley:

«La teoria lunar em fa mal de cap, i em produeix tan sovint insomni que desitjaria deixar-hi de pensar.»[2]

Algunes referències on s'aborden aquests problemes són [5], [20], [3], [12] i [13]. No és la meva intenció entrar en el contingut d'aquestes obres ni en els resultats concrets que s'hi demostren. El que voldria destacar és l'element que tenen en comú. En totes es proposa per a les observacions un model del tipus

$$Y = Xb + \varepsilon, \tag{1}$$

on

- $Y = (Y_1, \dots, Y_N)$ és el vector de les observacions. És a dir, són N observacions que es fan del fenomen que s'estudia, sota condicions idèntiques.
- $X = (x_{ij})$ és una matriu, $1 \le i \le N$, $1 \le j \le r$, $N \ge r$, coneguda.
- $b = (b_1, \dots, b_r)$ és un vector r-dimensional desconegut.
- $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)$ és un vector desconegut.

L'origen del model (1) és el següent: les lleis de la mecànica newtoniana, les fórmules de la trigonometria esfèrica, etc., depenent del problema que s'estudiï, postulen que allò que volem mesurar depèn de forma lineal de determinades variables $X_1, \dots, X_r$, amb coeficients $b_1, \dots, b_r$ que no coneixem. Si fixem un valor per a cadascuna de les variables $X_1, \dots, X_r$, per exemple $x_{11}, \dots, x_{1r}$, obtindrem per a la nostra incògnita

$$b_1 x_{11} + \cdots + b_r x_{1r}. \tag{2}$$

Això hauria de coincidir amb l'observació y_1. Però en general no serà així, ja que a la combinació lineal (2) s'hi afegeix una component d'error. És a dir,

$$y_1 = b_1 x_{11} + \cdots + b_r x_{1r} + \varepsilon_1 .$$

En repetir l'assignació de valors a les variables $X_1, \ldots, X_r$ N vegades (amb N prou gran perquè $N \geq r$), obtenim el model descrit a (1). Malgrat que ε és desconegut, com que representa un error, és raonable suposar que les seves components són molt petites.

La part matemàtica dels treballs que hem citat més amunt té com a objectiu en tots els casos determinar quina és la manera més efectiva de combinar els valors observats per proporsar-ne un com a *valor veritable* del que estem mesurant. En el llenguatge de la inferència estadística diríem que, a partir de les observacions $y_1, \ldots, y_N$, es tracta de fer una estimació de $b_1, \ldots, b_r$, que denotarem per $\hat{b} = (\hat{b}_1, \ldots, \hat{b}_r)$, i aleshores proposar

$$y = X\hat{b}$$

com a valor de l'amidament.

A [5] hi ha una proposta molt complexa de combinar les observacions. A [20] hi ha endemés una estimació dels errors comesos, treball que es completa a [12]. Potser el més interessant de tots és [3], on es desenvolupa un enginyós mètode geomètric que consisteix a trobar b de manera que ε compleixi:

(i) $\sum_{i=1}^{N} \varepsilon_i = 0$ (compensació dels errors),

(ii) la suma $\sum_{i=1}^{N} |\varepsilon_i|$ ha de ser mínima.

A [13] es dóna una justificació analítica del mètode proposat per Boscovich i Maire.

2. El mètode dels mínims quadrats de Legendre

Adrien-Marie Legendre (1752-1833), en un apèndix del treball *Nouvelles méthodes pour la détermination des orbites des comètes* (1805), que titula «Sur la méthode des moindres carrés», dóna una solució elegant al problema de fer una estimació de b, seguint la idea de Boscovich i Maire.

Legendre va ser un matemàtic polifacètic. Les seves contribucions més importants i conegudes són sobre les integrals el·líptiques, la teoria dels nombres i la geometria. Va formar part de la comissió que el 1795 va mesurar l'arc de meridià des de Montjuïc a Dunquerque, que serviria per donar la definició de *metre*. És molt probable que a partir d'aquí sorgís el seu interès pel problema que estem considerant.

La proposta de Legendre consisteix a fer una estimació de b imposant que $\sum_{i=1}^{N} \varepsilon_i^2$ sigui mínim. El problema té una solució geomètrica molt senzilla que exposem tot seguit. Suposem que la matriu X té rang r. Considerem el subespai vectorial de $\mathbb{R}^N$ generat per les columnes de X, $\mathcal{L} = \{X\, b,\, b \in \mathbb{R}^r\}$. Aleshores es tracta de trobar l'element de $\mathcal{L}$, $X\hat{b}$, tal que $\| Y - X\hat{b} \|^2$ sigui mínim. És conegut que

$$X\hat{b} = \mathrm{proj}_{\mathcal{L}} Y .$$

Així doncs, $Y - X\hat{b}$ haurà de ser ortogonal a $X\, b$, per a tot $b \in \mathbb{R}^r$, i això implica que

$$\hat{b} = (X^t X)^{-1} X^t Y .$$

Considerem un exemple, proposat per Legendre en el treball que hem citat, que permet de donar una interpretació física del seu mètode.

Es vol determinar la posició d'un punt en l'espai (y_1, y_2, y_3). En M amidaments diferents s'obté $y_1^1, \ldots, y_M^1$ per a la primera coordenada, $y_1^2, \ldots, y_M^2$ per a la segona, i $y_1^3, \ldots, y_M^3$ per a la tercera. El model lineal associat és $Y = X\, b + \varepsilon$ amb $Y = (y_1^1, \ldots, y_M^1, y_1^2, \ldots, y_M^2, y_1^3, \ldots, y_M^3)$, $X = Id$ (la matriu identitat en dimensió $N = 3\, M$) i $b = (y^1, \overset{M}{\ldots}, y^1, y^2, \overset{M}{\ldots}, y^2, y^3, \overset{M}{\ldots}, y^3)$.

És molt fàcil veure que la funció de les variables y^1, y^2, y^3,

$$\sum_{j=1}^{M} (y_i^1 - y^1)^2 + \sum_{i=1}^{M} (y_i^2 - y^2)^2 + \sum_{i=1}^{M} (y_i^3 - y^3)^2 ,$$

pren el seu valor mínim quan

$$y^1 = \frac{\sum_{i=1}^{M} y_i^1}{M}, \quad y^2 = \frac{\sum_{i=1}^{M} y_i^2}{M}, \quad y^3 = \frac{\sum_{i=1}^{M} y_i^3}{M}.$$

És a dir, l'estimació del punt (y^1, y^2, y^3) pel mètode dels mínims quadrats correspon a prendre el centre de gravetat de M masses d'intensitat $\frac{1}{M}$ situades en els punts (y_i^1, y_i^2, y_i^3), $i = 1, \ldots, M$.

Retrobem en aquest exemple una de les maneres tradicionals de combinar les observacions: les mitjanes aritmètiques.

Legendre feia una valoració molt positiva del seu mètode:

... *«L'application que nous allons faire de cette méthode à la mesure de la méridienne achévera de mettre dans tout son jour sa simplicité et sa fécondité.»* [17]

Si analitzem el treball de Legendre sota l'òptica que el que es resol és un problema d'interpolació, no podem fer-hi cap objecció. Es determina b imposant una condició molt raonable sobre l'error i, en conseqüència, fixats uns nous valors de $X_1, \ldots, X_r$, podrem *predir* el valor de y. En canvi, des del punt de vista d'un treball d'aproximació, Tobias Mayer, 50 anys abans, havia anat més lluny. La fiabilitat de la solució que Legendre proposa, *sa fécondité*, dependrà de si l'error que es fa en proposar $X\hat{b}$ com a síntesi de les observacions és prou petit. Però, com mesurar aquest error?

> «... *Il entre nécéssairement de l'arbitraire dans la distribution des erreus.*» [17]

Avui que disposem d'una teoria matemàtica que ens permet de mesurar la incertesa, la teoria de la probabilitat, podem dir que paradoxalment aquesta observació del mateix Legendre en dóna la solució. Legendre, però, no la tenia, la solució. Potser no havia copsat la importància d'aquella discussió entre la filosofia i la matemàtica que, alimentada pels jocs d'atzar, havia seduït altres matemàtics contemporanis, com, per exemple, Pierre Simon Laplace (1749-1827) i Karl Friedrich Gauss (1777-1855).

Els treballs de Laplace i de Gauss sobre aquest tema constitueixen una combinació perfecta de creativitat i profunditat matemàtica. El producte final pot sintetitzar-se en un resultat: el teorema del límit central. Arribem així a una etapa molt important en la història de la probabilitat. Aquest resultat representava la culminació en la construcció d'un cos teòric amb entitat pròpia (vegeu per exemple [16]), permetia un tractament matemàtic rigorós dels errors de les observacions i obria un camp important de recerca.

Per tal d'entendre millor les aportacions de Laplace i de Gauss retrocedirem quasi un segle i explicarem breument quins havien estat els intents de *quantificar la incertesa.*

3. Mesurar la incertesa: primeres aportacions

L'any 1654 va néixer a Basilea Jacob Bernoulli, d'una nissaga que va donar lloc a físics i matemàtics molt il·lustres. A partir de la descripció que en fa el seu primer biògraf ([6]), s'arriba a la conclusió que Jacob va ser més aviat desgraciat. El seu pare volia decantar-lo cap a la religió, i amb aquest objectiu va enfocar la seva formació cap al llatí, el grec i la filosofia. Un dia li va caure a les mans un tractat de geometria, i diuen que va quedar astorat davant la bellesa de les figures allí representades. Va decidir unilateralment canviar l'orientació dels seus estudis i aprendre matemàtiques i astronomia, com el seu germà Johannes, amb qui, val a dir-ho, les relacions van ser sempre dolentes. Al llarg de la seva vida va publicar molt poc. Fou Nicholas Bernoulli, el seu nebot, qui després de la seva mort li va recopilar l'obra probabilista. Aquesta va aparèixer publicada l'any 1713 en un llibre amb el títol d'*Ars Conjectandi.*

El resultat que m'interessa descriure de l'Ars Conjectandi és en la «Pars Quarta» i és el que avui coneixem com la llei dels grans nombres. El problema que motiva Bernoulli és de caire més aviat filosòfic: l'assignació empírica de probabilitats a esdeveniments, o assignació *a posteriori.*

Considerem una experiència aleatòria amb dos resultats possibles: A i A^c, per exemple, el llançament d'una moneda. Com determinar les probabilitats p i $1-p$ d'aquests resultats? La idea de Bernoulli consisteix a repetir l'experiència sota condicions idèntiques n vegades, i a comptar el nombre de vegades que hem obtingut A. Dividim aquest nombre per n i obtenim la freqüència relativa de l'esdeveniment A, $f_n(A)$. Amb arguments complicats demostra que, fixat $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left\{ \left| f_n(A) - p \right| > \varepsilon \right\} = 0 .$$

Aquesta és la primera versió demostrada de la llei del grans nombres. És un resultat brillant, però Bernoulli no n'estava prou content. Més aviat hauria volgut resoldre un problema plantejat en els següents termes, i que cau més directament en un context de mesurar la incertesa:

Fixats $\varepsilon > 0$ i $c > 0$, trobar $n \in \mathbb{N}$ tal que

$$P\left\{ \left| f_n(A) - p \right| > \varepsilon \right\} < c .$$

Bernoulli va resoldre aquesta qüestió d'una manera no gaire satisfactòria: els valors que troba per a n són massa grans. L'*Ars Conjectandi* acaba així bruscament. Una clara constatació de la seva decepció.

Siguin $\{ X_i, i = 1, \ldots, n \}$ variables aleatòries que prenen el valor 1, si A s'ha produït, i 0, si A no s'ha produït. La distribució de la variable aleatòria $\sum_{i=1}^n X_i$ és una binomial $B(n, p)$ i $f_n(A) = \frac{\sum_{i=1}^n X_i}{n}$. En la resolució del problema que acabem d'esmentar, Bernoulli es va veure confrontat amb càlculs difícils sobre la distribució binomial, i aquesta és la causa del fracàs aparent.

L'any 1718 Abraham de Moivre (1667-1754) publica *The Doctrine of Chances*. El resultat més destacable d'aquesta obra és una versió de l'avui anomenat teorema del límit central, o aproximació de la llei $B\left(n, \frac{1}{2}\right)$, convenientment normalitzada, per la llei normal. Concretament, utilitzant l'aproximació $n! \sim n^n e^{-n} (2\pi n)^{1/2}$, demostra que, per a $a, b \in \mathbb{R}$,

$$P\left\{ a \leq \frac{\sum_{i=1}^n X_i - \frac{n}{2}}{\left(\frac{n}{4}\right)^{1/2}} \leq b \right\} \xrightarrow[n\to\infty]{} \int_a^b \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx . \tag{3}$$

Dóna també un mètode per calcular la integral de la dreta, desenvolupant en sèrie l'exponencial i integrant terme a terme. Apareix en aquest resultat la densitat de la *llei normal*

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{x^2}{2} \right\} ,$$

que a partir d'ara tindrà un paper fonamental.

El resultat expressat a (3) permet de calcular la probabilitat del conjunt $\left\{ \left| f_n(A) - p \right| > \varepsilon \right\}$, de forma aproximada, en funció de n. Imposant que aquesta probabilitat sigui c, es pot trobar el valor de n, que surt sensiblement millor que l'obtingut per Bernoulli.

Aparquem momentàniament aquest resultat per reprendre el fil del nostre discurs sobre el model lineal que havíem deixat amb Legendre.

4. La corba dels errors

Per corba d'errors entenem una densitat de probabilitat sobre $\mathbb{R}$ segons la qual suposem que es distribueixen els errors. En el marc del model lineal i amb el llenguatge actual, diríem que ε és una variable aleatòria amb distribució absolutament contínua, la seva densitat és la *corba d'errors*. Designarem a partir d'ara amb $\phi(x)$ aquesta corba.

Al llarg del segle XVIII hi va haver diverses aportacions sobre aquest tema. Destaquem les de Simpson (1755), Lambert (1760), Lagrange (1769) i Laplace (1774, 1781). Segurament la segona cita de Legendre que hem fet a l'apartat 2 es refereix a aquesta corba. Cal, però, insistir que no es relacionava, clarament la corba dels errors amb el problema de com combinar les observacions, i que és a la base de la inferència estadística.

Sembla raonable imposar que una corba d'errors $\phi(x)$ compleixi les propietats següents:

(i) $\phi(x) = \phi(-x)$. En efecte, no hi ha raons objectives per suposar més probables els errors positius que els negatius.
(ii) $\phi(x) = 0$ si $x \notin [-a, a]$, ja que els errors no poden ser arbitràriament grans.

Simpson, per exemple, proposa una distribució triangular fent també el càlcul de la distribució de la mitjana empírica dels errors.

L'aportació de Laplace té un caràcter més analític. Com a bon newtonià, pensava que la corba dels errors s'havia de poder derivar matemàticament a partir de certs principis que calia especificar. Aquests són: ϕ ha de ser una funció parell, per la raó que hem donat abans; ϕ com a funció d'x ha de ser *decreixent*, perquè els errors grans són menys probables; finalment, una condició sobre el decreixement. Argumenta que no hi ha motius per suposar un comportament diferent de les ordenades i dels seus increments, condició que formula així:

$$\frac{d\,\phi(x)}{dx} = -m\,\phi(x) \ . \tag{4}$$

La solució d'aquesta equació diferencial, imposant que $\phi(x) = \phi(-x)$ i que $\int_{\mathbb{R}} \phi(x)dx = 1$ (recordeu que ϕ ha de ser una densitat de probabilitat), és

$$\phi(x) = \frac{m}{2}\, e^{-m(x)} \ . \tag{5}$$

Observeu que $\lim\limits_{(x)\to\infty} \phi(x) = 0$; aquesta propietat és lleugerament diferent de l'expressada en (ii), però igualment raonable per un error. Des del punt de vista matemàtic, la deducció de la corba d'error donada per (4) era impecable. Ara bé, una anàlisis estadística basada en aquesta densitat no és senzilla, i endemés el paràmetre m és desconegut.

En síntesi, els ingredients presentats fins ara ens porten a les consideracions següents: la natura estocàstica dels errors es transmet a les observacions; d'altra banda, tenim un mètode, el dels mínims quadrats, per combinar aquestes observacions aleatòries; tenim també elements d'una teoria per quantificar l'atzar. És, doncs, natural plantjar-se la qüestió següent: com es pot, mitjançant la probabilitat, mesurar la bondat d'ajustament del mètode dels mínims quadrats? Per donar-hi una resposta calia formular el problema de la combinació d'observacions des d'una perspectiva probabilista. Aquesta va ser la gran innovació de Gauss.

5. El model lineal segons Laplace i Gauss

L'any 1809 Gauss publica la *Theorie motus Corpurum Celestium*... [7], on es fa una investigació matemàtica de les òrbites dels planetes. L'última secció d'aquest llibre es dedica al problema que ens ocupa: com combinar les observacions. Gauss presenta el que avui anomenem model lineal de regressió múltiple, és a dir, el model (1), amb les hipòtesis especificades a l'apartat 1. Imposa endemés la condició següent:

- Les variables aleatòries $\varepsilon_1, \dots, \varepsilon_N$ són independents i tenen una densitat donada per una corba d'errors $\phi(x)$ (que compleix (i) i (ii) de l'apartat anterior), tal que presenta el seu valor màxim en $x = 0$.

La idea de Gauss pot resumir-se així: si hem observat els valors $y_1, \dots, y_N$ és perquè aquests són els més probables; dit altrament, els errors comesos, $\varepsilon_1, \dots, \varepsilon_N$, són els més probables. Per tant, l'estimació de b es farà de manera que la funció

$$\phi(\varepsilon_1). \dots .\phi(\varepsilon_N) = \phi(y_1 - (X\,b)_1). \dots .\phi(y_N - (X\,b)_N) \tag{5}$$

sigui màxima.

El primer membre de la igualtat (5) és la funció de versemblança del model estadístic definit per (1); el mètode d'estimació proposat per Gauss és el del màxim de versemblança.

Per tal de trobar el valor màxim de $\prod_{i=1}^{N} \phi(\varepsilon_i)$ cal conèixer la funció ϕ. Gauss procedeix de la manera següent: imposa que ϕ sigui tal que el valor més probable d'una variable observada sota condicions idèntiques correspongui a la mitjana aritmètica i, en aquest cas particular, o sigui quan el model és

$$y_1 = b + \varepsilon_1, \dots, y_N = b + \varepsilon_N.$$

demostra que $\hat{b} = \frac{\sum_{i=1}^{N} y_i}{N}$ maximitza la funció $\prod_{i=1}^{N} \phi(\varepsilon_i)$ únicament si $\phi(x)$ és la densitat d'una llei $N(0, \sigma^2)$, on $\sigma > 0$ és un paràmetre desconegut.

Aquest resultat li suggereix completar les hipòtesis sobre el model: imposa que la llei de les variables aleatòries $\varepsilon_1, \ldots, \varepsilon_N$ sigui una $N(0, \sigma_2)$. A continuació constata que la versemblança és màxima pel valor b obtingut pel mètode dels mínims quadrats de Legendre. En efecte, sota la hipòtesi de normalitat el segon membre de la igualtat (5) val

$$\frac{1}{(\sqrt{2\pi}\,\sigma)^n} \exp\left\{ -\frac{\|Y - Xb\|^2}{2\sigma^2} \right\}$$

Com que l'exponencial negativa és una funció decreixent del seu argument, maximitzar (5) equival a minimitzar $\| Y - Xb \|^2$.

Resumint, cal destacar tres punts de l'argument de Gauss:

(1) La mitjana aritmètica, obtinguda en models particulars pel mètode dels mínims quadrats de Legendre, és el valor més probable únicament si els errors segueixen una llei normal.
(2) Com que la mitjana aritmètica ha estat una manera habitual de combinar les observacions, sembla justificat prendre la distribució normal com a corba d'errors.
(3) Sota la hipòtesi de normalitat, el mètode de Gauss dóna el mateix estimador que el mètode dels mínims quadrats.

Gauss completa així el programa iniciat per Legendre utilitzant idees innovadores, malgrat una certa manca de rigor en la justificació de la hipòtesi de normalitat. Analitzem ara l'aportació de Laplace.

L'any 1805 Laplace acaba el quart volum de l'obra *Mécanique céleste* i torna a interessar-se per problemes de la probabilitat. El 1810 llegeix a l'Acadèmia de Ciències de París una memòria sobre una extensió del teorema del límit central de Moivre per a variables discretes.

L'única relació entre el problema del límit central i la determinació de la corba dels errors en l'obra de Laplace fins al 1810 era *metodològica*. La determinació de la corba dels errors el porta, en una anàlisi posterior, a la determinació de la distribució de la mitjana aritmètica dels errors. Amb aquest objectiu desenvolupa una tècnica que ha resultat molt útil en el tractament de la independència: el que avui coneixem per *funció generatriu* (per a distribucions discretes) i la *transformada de Laplace* (per a distribucions contínues).

Laplace va conèixer els resultats de Gauss poc després de llegir a l'Acadèmia de Ciències el seu treball sobre el teorema del límit central. Va quedar totalment impressionat en descobrir un nexe entre la corba dels errors i el resultat contingut en el teorema del límit central que servia per presentar el mètode de Gauss amb menys fissures. En efecte, els errors poden considerar-se com una suma d'ingredients aleatoris i, segons el teorema del límit central, es pot suposar que es distribueixen aproximadament segons una llei normal. En definitiva, la coherència entre el mètode de Gauss i el dels mínims quadrats s'estableix a partir d'uns principis científics. La filosofia newtoniana no es posa en perill.

Aquesta aportació de Laplace es recull en la publicació del teorema del límit central l'any 1811. A partir d'aquesta data, i fins al 1823, Laplace i Gauss, en un meravellós *tour de force*, completen l'anàlisi estadística del model lineal de regressió múltiple: estimació de σ, distribucions dels estimadors obtinguts, intervals de confiança i prova de la propietat que assegura que entre els estimadors lineals de b, el dels mínims quadrats és el que dóna variància (error) mínima.

6. Epíleg

La història del model lineal i les seves ramificacions no es clou amb Laplace i Gauss. Altres científics, com per exemple Bienaymé i Cauchy, continuen interessats en el tema i en fan aportacions. I arribem a Paul Lévy (1886-1971).

...«*Pour la théorie des erreurs je me rappelais seulement que les erreurs accidentelles obéissent à la loi de Gauss, et que c'était dû à ce qu'elles résultent de diverses causes indépendantes, dont chacune a un effet négligeable.*» [18]

Aquesta cita es refereix a un record sobre el curs que P. Lévy va seguir de H. Poincaré a l'École Polytechnique de París. Recull un element essencial: la justificació de Laplace del mètode proposat per Gauss. Però aquesta justificació havia estat demostrada rigorosament només per a variables aleatòries discretes, i per tant era incompleta.

L'any 1919 el director de l'École Polytechnique va proposar a P. Lévy de fer tres conferències sobre el càlcul de probabilitats i la teoria dels errors. Havia passat un segle des de la solució de Laplace-Gauss. Els mètodes de l'anàlisi matemàtica havien impregnat la probabilitat, i l'exigència en el nivell de rigor no era la mateixa que la de 1811.

«*Je n'admis pas cette démission de la science. Si c'est vrai que les erreurs obéissent à la loi de Gauss on doit pouvoir l'expliquer, et si ce n'est pas fait, je le ferai.*» [18]

D'aquesta manera, l'estudi del model lineal, ja gairebé complet, dóna pas a l'anàlisi del problema del límit central amb tota generalitat. Paul Lévy, juntament amb altres probabilistes de l'escola de Sant Peterburg i d'Estocolm, en varen fer les contribucions més relevants, creant els elements teòrics que dotaren la teoria de la probabilitat d'entitat pròpia com a branca de la matemàtica.

Llafranc, Baix Empordà
Agost de 1990 i 1991

Referències

[1] BERNOULLI, Jacob. (1713), *Ars Conjectandi*. Basel: Thurnisiorum.

[2] BERRY, A. (1898), *A Short History of Astronomy*. Reprinted, 1961; New York: Dover.

[3] BOSCOVICH, Roger Joseph and Christopher Maire. (1755), *De Litteraria Expeditione per Pontificiam ditionem ad dimetiendas duas Meridiani gradus*. Rom: Palladis.

[4] DE MOIVRE, Abraham. (1718), *The Doctrine of Chances*. London: W. Pearson.

[5] EULER, Leonhard. (1749), "Recherches sur la question des inegalités du mouvement de Saturne et de Jupiter". Reprinted, 1960 in Leonhardi Euliri, Opera Omnia, 2nd series, vol. 25, 45-157. Basel: Turici.

[6] FONTENELLE, Bernard. (1717), *The Lives of the French, Italian and German Philosphers, Late Members of the Royal Academy of Sciences in Paris*. London: W. Innys.

[7] GAUSS, Karl Friedrich (1809), *Theorie motus corporum celestium*. Reprinted, 1963; New York: Dover.

[8] HEYDE, C.C. and SENETA, E. (1977), *I.J. Bienaymé, Statistical Theory anticipated*. Studies in the History of mathematics and Physical Sciences 3. New York Heidelberg, Berlin: Springer-Verlag.

[9] LAGRANGE, J.L. (1766), "Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de pluiseurs observations". Reprinted in Lagrange, 1867-1892, vol. 2, 173-236.

[10] LAPLACE, P.S. (1774), "Mémoire sur la probabilité des causes par les événements". Reprinted in Laplace, 1878-1912, vol. 8, 27-65.

[11] — (1781), "Mémoire sur les probabilités". Reprinted in Laplace, 1878-1912, vol. 9, 383-485.

[12] — (1788), "Théorie de Jupiter et Saturne". Reprinted de Laplace, 1878-1912, vol. 11, 95-239.

[13] — (1793), "Sur quelques points du systéme du monde". Reprinted in Laplace, 1878-1912, vol. 11, 477-558.

[14] — (1799-1805), *Traité de mécanique céleste*. Reprinted in Laplace, 1878-1912, vols. 1-4.

[15] — (1811), "Mémoire sur les intégrales définies et leur application aux probabilités". Reprinted in Laplace, 1878-1912, vol. 12, 357-412.

[16] — (1812), *Théorie analytique des probabilités*. Reprinted in Laplace, 1878-1912, vol. 7.

[17] LEGRENDE, A.M. (1805), *Nouvelles méthodes pour la détermination des orbites des comètes*. Reprinted, 1959, New York: Dover.

[18] LÉVY, P. (1970), *Quelques aspects de la pensée d'un Mathématicien*, Paris: Librairie Scientifique et Technique Albert Blanchard.

[19] MAISTROV, I.E. (1974), *Probability Theory, a historical sketch*. New York and London: Academic Press.

[20] MAYER, Tobias (1750), "Abhandlung über die Umwalzung des Monds um sie-

ne Axe und die schienbare Bewegung der Monds flechten". *Kosmographische nachrichten und Sammlungen auf das Jahr* 1748, 52-183.

[21] SIMPSON, T. (1755), "On the advantage of taking the mean of a number of observations in practical astronomy". *Philosophical Transactions of the Royal Society of London*, **49**, 82-93.

[22] STIGLER, S.M. (1986), *The History of Statistics. The Measurement of Uncertainty before 1900.* Cambridge, Massachusets, and London, England: The Belknap.